DOCUMENT RESUME

ED 433 173 RC 022 081

AUTHOR	Howley, Craig B.
TITLE	The Matthew Project: State Report for Montana.
INSTITUTION	Ohio State Univ., Columbus.; Appalachia Educational Lab., Charleston, WV.
SPONS AGENCY	Rural Challenge Policy Program, Randolph, VT.
PUB DATE	1999-08-08
NOTE	53p.; For related reports, see RC 022 082-083 and RC 022 087-088.
PUB TYPE	Numerical/Quantitative Data (110) -- Reports - Research (143)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Academic Achievement; American Indians; Correlation; Economically Disadvantaged; Elementary Secondary Education; Enrollment; Regression (Statistics); *School District Size; *School Size; *Small Schools; *Socioeconomic Influences; *Socioeconomic Status; Tables (Data)
IDENTIFIERS	*Montana; Small School Districts

ABSTRACT

Previous studies found that the small size of schools or school districts mitigated the negative influence of poverty on academic achievement in California, Alaska, and West Virginia. The Matthew Project extends this research in four additional states selected to provide varied settings. Montana has a remarkable number of small schools and small districts, has a substantial minority (American Indian) population, and has not yet joined the systemic reform craze sweeping the nation. In this report, Montana data were used in regression equations that predict overall school or district achievement from measures of size, socioeconomic status (SES), and the product of size and SES. These equations illuminate possible "excellence effects" of size by showing which communities (based on SES-level) may benefit or lose from increases in school or district size. Equity effects of size on achievement were also tested by computing the correlation between SES and achievement in groups of larger and smaller schools and districts. The excellence effect was comparatively weak, but the equity effect of size was very strong, with dramatically reduced correlations between SES and achievement for smaller schools and districts. Montana leads the nation in National Assessment of Educational Progress scores, while spending average per-pupil amounts on public schooling. Montana's small-scale schooling probably cultivates academic achievement in impoverished communities, a conclusion actually supported by the weak excellence effect. Appendices include a regression model and statistical data. (Contains 21 references and 14 data tables.) (SV)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 433 173

The Matthew Project:
State Report for Montana

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Craig Howley

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Craig B. Howley, co-investigator
Ohio University and Appalachia Educational Laboratory
75619 Lively Ridge Road, Albany, OH 45710
howleyc@ael.org

BEST COPY AVAILABLE

August 8, 1999

RC022081

The Matthew Project: State Report for Montana

EXECUTIVE SUMMARY

The Matthew Project, with support from the Rural Challenge Policy Program, investigated the possible academic excellence and equity effects of school and district size in Montana, Georgia, Ohio, and Texas. Previous studies (Friedkin & Necochea, 1988; Howley, 1995, 1996; Huang & Howley, 1993) had reported that the effects of size depended on accounting for the interaction of size and socioeconomic status (SES). The Matthew Project studies build on that line of inquiry.

We constructed equations that predict overall school or district test scores from measures of size, socioeconomic status, and the product of size and socioeconomic status. Appendix A describes the methodology in detail. These (regression) equations provide a view of the possible excellence effects of size because they show which SES levels are likely to benefit from, or conversely, to be harmed from increases in school or district size, and to what extent. Good policy should aim to distribute resources in such a way that students of various SES levels are benefitted to the maximum; such arrangements would, in effect, cultivate a degree of excellence by providing conditions that appear to maximize academic achievement.

We also tested the equity effects of size (on academic achievement) by dividing relevant groups of schools and districts into two equal groups at the median of size. Then we computed the correlation between SES and achievement. Weaker correlations indicate that SES accounts

for less of the variability in achievement, and so indicate a weakening of the bond between these two qualities. Equity in achievement depends on the weakening of this bond; for instance, on *disrupting* or *mitigating* the negative influence of poverty on achievement. This goal reflects that view, for instance, that achievement should be more strongly influenced by such qualities as effort, adequacy of educational funding, and fairly distributed opportunities to learn.

Montana operates many small schools and districts. In fact, on average, each Montana district would be seen to operate just two schools. This average, however, would obscure the fact that many of Montana's districts are one-school districts. Further--a fact that would likely astound educators and citizens outside Montana--75% of this state's elementary school *and* high schools enroll fewer than 300 students. These facts makes Montana a radical example of small-scale schooling; by not consolidating all its schools and districts in accord with the conventional professional wisdom that used to apply (i.e., during most of the previous century) to school size matters, Montana finds itself ahead of the rest of the nation so far as the new conventional wisdom is concerned. This new conventional wisdom insists that the optimum school sizes are much smaller than previously thought. The new conventional wisdom, however, remains silent about district size. The Matthew Project questions the appropriateness of views of optimum size, first, and also investigates the possible effects of district size--all in relation to achievement.

Our variables were size (computed as grade cohort enrollment or, for some district-level analyses, total enrollment), rates of free-and-reduced-price meals, the product of the two (our "interaction term"), and total school- or district-level achievement (average of 5 subject areas) for grade 4, grade 8, and grade 11. For some district analyses we also used the student-weighted average of all test grade-level scores.

Evidence for an excellence effect exists in Montana, but it is comparatively weak and is most clearly evident at the school level for grade 4; some evidence points to a grade 8 effect at the school level. Evidence for an excellence effect is clearest at the district level only among Montana's 51 K-12 districts. The rest of Montana's 457 districts, however, are either elementary or high school districts.

Evidence for an equity effect of sized on academic achievement, by comparison, is very strong. Practically all comparisons at the school and district level show dramatically reduced correlations (which indicate even more dramatically reduced explained variance, which is the square of the correlation coefficient) for smaller versus larger groups. In some cases, the association between SES and achievement is not statistically different from 0 among the smaller groups. In the near absence of an excellence effect, the persistence of an equity effect of school and district size is striking, and even surprising.

Finally, other evidence presented in this report shows that (1) Montana leads the nation in scores on the National Assessment of Educational Progress, (2) Montana teachers are not, on average, paid extremely well, (3) Montana spends no more than an average per-pupil amount to support public schooling, (4) as a proportion of their incomes, Montanans appropriate more than other states to support public education, and (5) among all the states, income distribution is comparatively equal in Montana.

Montana appears to have derived substantial benefits from its historic decision to maintain small schools and districts. Wise policy would necessarily balance achievement of excellence, equity, and cost. The NAEP results show the excellence results, but the prevalence of small schools also suggests that Montana's small-scale schooling probably works well to

cultivate academic achievement in impoverished communities, which, in other states, do not enjoy the mitigating effects of small size. The weak "excellence effect" actually attests to this conclusion.

The Matthew Project

The Matthew Project, with funding from the Rural Challenge Policy Program (now known as the Rural School and Community Trust Policy Program), investigated the possible academic excellence and equity effects of school and district size in Georgia, Montana, Ohio, and Texas. The project title refers to a parable about stewardship in the gospel according to Matthew (13:12): "For whosoever hath, to him shall be given, and he shall have more abundance: but whosoever hath not, from him shall be taken away even that he hath." Building on previous research efforts in Alaska, California, and West Virginia, the Matthew Project was particularly concerned to investigate the possible contributions of smaller school size to academic success in impoverished communities.

BEST COPY AVAILABLE

The Matthew Project: State Report for Montana

Poverty figures as the chief and most prevalent threat to normal academic accomplishment among individuals. If your family is poor, your own odds of succeeding in school lengthen. Your odds are longer still if you attend school with children from many other poor families, which is a likelihood in the U.S., since schools are segregated by social class[1]. In any case, it is certain that affluent communities enjoy decent schools and high-minded pedagogy, whereas impoverished communities continue to "enjoy" shabby schools and a pedagogy of expedience (that is, schooling that is primarily custodial). In fact, one might say that as the threats increase among increasingly impoverished communities, the local resources to counter them diminish *simultaneously*. It seems an especially vicious arrangement.

If, however, some quite ordinary and easily appreciated feature of schooling could be so deployed as to resolve this dilemma more favorably for the children of society's least privileged members, we should applaud it and move to deploy it as seemed advisable. The Matthew Project has pursued a promising line of inquiry relevant to such a hope.

The Matthew Project

The Matthew Project [2] extends previous investigations that found small size to mitigate

---

[1]San Francisco recently adopted a plan to integrate its schools on a socioeconomic instead of a racial basis. "Ironically, the fact that economic segregation was never found unconstitutional means that voluntary measures addressed to class are constitutionally permissible" (Kahlenberg, 1999, p. 30-52).

[2]"The Matthew Principle" (Howley, 1995) is the title of the article reporting a study that the current work extends; that previous article is available on the Web at the following URL: <http://olam.ed.asu.edu/epaa/v3n18.html>.

the negative influence of poverty on achievement (in California, Alaska, and West Virginia). In general, the implication from such findings is that the more impoverished the community, the smaller the school or district that would be required to maximize overall student achievement.[3]

This line of inquiry tests the "interaction hypothesis" of school and district size. The interaction hypothesis expresses the possibility that the degree (i.e., strength or weakness) and directionality (positive or negative) of the relationship of size to achievement is contingent on community SES. That is, no one size is "best" or "optimal," because the effects of size hypothetically vary among communities with differing levels of SES.

This is an important insight because so many previous studies have found that the relationship between size and achievement is, in general, small if not insignificant. That research, however, has measured the relationship of size and achievement across all schools and districts. By contrast, the interaction hypothesis suggests that in some places the relationship could be negative and in some places positive; and that, in some places it could be weak, and in some places strong. Further, as a formal and testable hypothesis, it says that this variation could be *systematically associated* with changes some other condition. What might such a condition be? Socioeconomic status (SES) is well known to be the strongest single influence on student achievement, and so it is a logical choice for our "contingent condition."

In informal terms, then, the Matthew Project asks whether or not size is beneficial or harmful to students, and to what extent, in view of community socioeconomic status. Previous

---

[3]In state after state, legislatures and State Education Agencies (SEAs) have, in myriad laws, regulations, and pronouncements, insisted that teaching students to high standards, and ensuring high levels of student achievement, be the first aim of schooling.

studies in California, Alaska, and West Virginia confirmed such an influence[4].

In the Matthew Project studies, we use school and district performance on state-mandated standardized tests as the measure of achievement. In some states these tests are norm-referenced and in some states they are criterion-referenced (for instance, some states require students to "pass" proficiency tests, and the percent passing in a school or district becomes a gauge of accountability). In any case, in the Matthew Project, schools and districts (not individual students) are the object of study (called "the unit of analysis" in the language of researchers).

The Matthew Project has extended the line of research to include investigations in four additional states: Georgia, Ohio, Montana, and Texas. We selected these states to provide a range of relevant settings to help test the prevalence of the *interaction hypothesis* (the key insight in this line of research) further--a group of states that exhibit variability in the conditions that affect decisions about school and district size. Aside from requiring the availability of suitable data in all states, our selection aimed to provide geographic variety, variety in school district organization (township and whole-county organization), ethnic and racial diversity (both within and between states), and rural and urban diversity (again, within and between states). Table 1 displays this diversity by state.

---

[4]Friedkin & Necochea, 1988; Howley, 1995, 1996; Huang & Howley, 1993.

Table 1
Variability in Matthew Project States (Illustrative Data)

	N students litigation (thousands)[1]	N schools[2]	N districts[2]	% minority students[2]	% rural schools[2]	district organization[3]	finance ruling
GA	1,100	1,817	180	44%	21%	county/mixed	constitutional
MT	152	889	457	13%	60%	independent	unconstitutional
OH	1,793	3,841	611	18%	28%	independent[4]	unconstitutional
TX	3,237	7,053	1,042	55%	24%	independent	unconstitutional

Notes: 1. data from Johnson (1989)
2. data from National Center for Education Statistics (1998)
3. Montana maintains elementary, secondary, and K-12 districts.
4. Several of Ohio's poorest districts have been consolidated into county units.

BEST COPY AVAILABLE

Table 2
Education Resource Allocations by State[1]

	GA	MT	OH	TX
% kids in elementary schools < 350	8	56	24	10
% kids in secondary schools < 900	17	57	49	24
% of schools with class size < 25	64	80	52	90
% schools in need of major repair	26	20	38	27
% **classrooms** with Internet access	35	55	50	42
% **hi pov schools** with net access	87	85	63	79
% all other schools with net access	93	80	81	83
per pupil spending	$4,595	$5,428	$5,438	$4,996
district pp spending disparity[2]	$1,628	$9,171	$5,804	$4,210
education spending ratio[3]	$38	$55	$40	$48
average teacher salary	$35,688	$30,604	$38,833	$35,148
Ed Week Equity Grade	B	B-	C+	D
Ed Week Adequacy Grade	B-	C	B	C+
Ed Week Accountability Grade[4]	A	D	A-	A
Ed Week School Climate Grade[5]	C	B-	D+	C+

Notes. 1. all data from Education Week (1997, 1998)
2. difference between per pupil spending of districts at the 95th and 5th percentiles on spending
3. education spending for every $1,000 of per capita income
4. Education Week "accountability grade" is based on the degree to which states adopt "high standards for all children and assessments aligned with those standards," a perspective that does not enjoy universal support.
5. School climate grades varied from B+ (one state, VT), B (one state, ME) and B- (three states) to D- (three states). 27 states earned grades of C-, C, or C+.

BEST COPY AVAILABLE

Table 2 presents salient data from comparisons that appeared in *Education Week* special supplements in 1997 and 1998, much of it concurrent with the data analyzed in the Matthew Project state-level reports. The state-level differences apparent in Table 1 are, if anything, magnified by the information in Table 2. Together, Tables 1 and 2 confirm the existence of substantial differences among these four states.

Finally, Table 3 reports widely accessible state-level aggregate scores on National Assessment of Educational Progress proficiency tests administered in 1996 (eighth-grade mathematics) and in 1998 (eighth-grade reading). The performance of students at this grade level is, in a sense, the acid test for school effectiveness (with achievement the touchstone), because their performance shows the accumulated effects of instruction within the state system but before the attrition of school leaving takes its largest toll in high school. Among the three states that participate in this testing program[5] (and, indeed, nationally) Montana has a substantial history of high achievement, as probed by NAEP.

---

[5]Ohio no longer participates in the NAEP state-level testing program. However, in the 1992 tests, 59% of Ohio 8th graders scored at or above the basic level in math (Education Week, 1997).

Table 3
National Assessment of Educational Progress Results

	Reading (1998)		Mathematics (1996)	
	% of students testing at or above "basic"	average NAEP score	% of students testing at or above "basic"	average NAEP score
GA	51%	257	51%	263
MT	83%	270	75%	283
OH	---	---	---	---
TX	76%	262	59%	270

Notes: Aggregate NAEP state-level reading scores range from about 230 to 270; aggregate state-level math scores range from about 233 to 284.

Clearly the educational systems differ markedly in these two states--they differ on measures of structure and governance, on measures of resource allocation (processes of the education system), and on outcomes in comparison to one another and to the nation as a whole, according, at least, to the National Assessment of Educational Progress, arguably among the most carefully designed of the various accountability assessment schemes. Those who insist on fashioning "nationally representative" pictures of educational processes will apparently miss a great deal of variation associated with between-state differences.

Educational systems themselves are nonetheless legitimately viewed as parts of larger social systems (e.g., Friedkin & Necochea, 1988; Weber, 1947), that is, in this case, as parts of the political economies of the various states. Most relevant to the Matthew Project, however, are measures of economic equity. Table 4 presents a picture of the prevailing diversity of economic equity in these four states.

BEST COPY AVAILABLE

Table 4
Selected Economic Equity Indicators by State[1]

	GA	MT	OH	TX
income distribution ratio (rank)[2]	10.9 (40)	7.9 (11)	9.9 (30)	12.2 (44)
rural-urban disparity score (rank)[3]	96 (8)	83 (1)	166 (31)	217 (48)
crime rate (rank)[4]	6301 (45)	4494 (22)	4456 (20)	5709 (38)
sectoral diversity score (rank)[5]	.0445 (11)	.0517 (17)	.0563 (23)	.0548 (21)
dynamic diversity scores (rank)	2.46 (32)	-1.54 (8)	1.71 (29)	2.11 (30)
CED 1998 equity subindex[6]	B (20)	B (13)	C (33)	F (49)

Notes. 1. all data from Clones (1998); numbers in parentheses represent state rankings (lowest number is *most favorable* rank)
2. ratio of earnings of highest income quintile to lowest (see Clones, 1998, pp. 76-77)
3. index of economic disparity between rural and urban areas; the higher the score, the greater the disparity (see Clones, 1998, pp. 76-77)
4. FBI index, serious crimes per 100K population (see Clones, 1998, pp. 78-79)
5. sectoral and dynamic diversity measure (a) diversity of traded industries and (b) similarity of changes in employment among the state's key traded industries (see Clones, 1996, pp. 84-5)
6. The equity subindex is one of three measures comprising the CED "Economic Performance Index"; there are three overall composite indices. The various state grades (GA, MT, OH, and TX) on these are as follows: (1) economic performance (A, C, C, C); (2) business vitality (A, D, C, C); and development capacity (C, B, A, C); see Clones, 1998, pp. 14-19.

BEST COPY AVAILABLE

Overall, in the view of the CED report (Clones, 1998), Montana exhibits the most economic equity and Texas the least among these four states. Nationally, Alaska ranks first and Texas last in the CED report. We do not, of course, trace the exact links between the political economic context in these reports. In any case, the equity and excellence of achievement are of central concern in these studies, and these concerns relate to the fiscal equity and adequacy of state education systems, and these qualities relate in turn quite clearly relate to conceptions of the manner and form of the distribution of economic goods in society at large. The 50 states vary on all these measures, and so do those examined in these reports.

The extent to which the interaction hypothesis of school and district size appears may well be influenced by such differences as indicated in Tables 1-4. In any case, we believe that studies that attempt to synthesize a national picture of education miss the point that the US maintains a dramatically decentralized system in which longstanding experimentation and local options have evolved quite different state-based systems of schooling. The results of these experiments should be of interest to those who make policy, those who teach, and those who vote. The Matthew Project reports bear witness, we believe, to the importance of state-based educational studies.

Montana Policy Context (or, Why Montana?)

Montana is among the 15 states[6] whose education finance schemes have been ruled

---

[6]In reverse chronological order: OH (DeRolph v. State, 1997); VT (Brigham v. State, 1996); AZ (Roosevelt Elementary v. Bishop, 1994); MA (McDuffy v. Secretary, Executive Office of Education, 1993); TN (Tennessee Small School Systems v. McWherter, 1993); NJ (Abbot v. Burker, 1990); KY (Rose v. Council for Better Education, 1989); TX (Edgewood v. Kirby, 1989); MT (Helena Elementary v. State, 1989); AR (DuPree v. Alma School District, 1983); WY (Washakie v. Herschler, 1980); WV (Pauley v. Kelley, 1979); WA (Seattle School District No. 1 v. State, 1978); CT (Horton v. Meskill, 1977); and CA (Serrano v. Priest, 1976).

unconstitutional by state courts (Helena Elementary v. State, 1989). At the time of the ruling a 1:9 ratio existed between the district spending the least per pupil and the district spending the most per pupil. As of this writing Montana guarantees each district a supplement sufficient to bring funding in each district up to a hypothetical state average (based on average statewide property valuation and assessment figures).

As Tables 1-3 suggest, however, Montana is, in comparison to the other states selected (and indeed, to the US as a whole) interesting for a number of other reasons that reflect its history and development to some degree. First, it has retained a remarkable number of small schools and an even more remarkable total number of school districts (n=457, most quite small). Table 1 suggests that, on average, the typical Montana school district would operate just two schools. In fact, many Montana districts operate just one school. The scale of the school enterprise is quite different in Montana than in other states. Additionally, Montana has, for a largely rural state, a substantial minority population--13 percent. Other very rural states, such as Vermont, Maine, and West Virginia have 4 percent or fewer students who are members of ethnic minorities. Finally, Montana has so far refused to join the systemic reform craze (i.e., the standards-assessment-accountability phenomenon), which has captured attention in SEAs across the nation. Rural-inspired equity suits preceded "systemic reform," so a connection is possible. Montana is experiencing some movement toward accountability, embarrassed, as it must be by the perception that it is out of step with the national trend.

In this light it is interesting to observe that a study sponsored by the Montana School Boards Association and led by former Montana Governor Ted Schwinden (Schwinden & Brannon, 1993) advised the state against sponsoring any wide-ranging program of school

---

The Kentucky ruling is unique in that it found the entire state system of public schooling--including school finance--unconstitutional.

consolidation. Schwinden and Brannon reviewed the literature on school size, replicated one of the few studies to test the claim that consolidation saved money, and concluded that little was to be gained by further consolidation in Montana. In many other states with numerous independent districts, consolidation is viewed in State Education Agencies (SEAs) either openly or tacitly as a problem. This may be the case, as well, in Montana, but, if so, public support for small-scale schooling and concomitant (possibly related) resistance to systemic reform has prevailed thus far.

Methods

The Matthew project has conducted a series of studies in which equations relate size of schools or districts, average socioeconomic status of those same schools or districts, and the interaction of size and socioeconomic status[7] in order to predict the aggregate student achievement[8] of schools and districts. That is, the performance of schools and districts--not individual students--was what we sought to predict. These equations all look something like this, and are really quite simple:

$$\text{size} + \text{SES} + (\text{size} \times \text{SES}) = \text{achievement}$$

If, in these equations, the interaction term proved statistically significant, we took that fact to mean that the influence of size on achievement varied systematically in tandem with SES. This being the case, we calculated the size of that effect (effect size) using a method pioneered by Friedkin and Necochea (1988) and applied subsequently in Howley (1995, 1996). The model and procedure is more fully explained in Appendix A. For all equations in this report we adopted

---

[7]independent variables

[8]dependent variable

Chronbach's (1987) method of centering independent variables in order to reduce the collinearity of related independent variables--a procedure particularly important in the case of our interaction term[9].

In some cases, we performed additional analyses in order to help draw out the practical implications of findings. One such analysis that we performed for every state concerns an equity effect of small size of schools and districts. These analyses we performed regardless of whether or not the interaction hypothesis was confirmed. One effect of small size, in view of the interaction hypothesis, is that smaller units mitigate the damaging effects of poverty on achievement. That is, "excellence" is more closely approximated when impoverished communities are served by small schools and districts. In this case, it would seem that the small size helps disrupt the usually strong relationship between SES and achievement. This means, that in impoverished communities, excellence is cultivated via an apparent equity effect (that is, breaking the usual bond between SES and achievement). The equity question is whether or not this phenomenon actually pertains to small schools across the board--*regardless of community SES*. In the Matthew Project we tested this possibility by dividing districts and schools at the median of size (the size that divides the small half from the large half) and computing the correlation (Pearson r) between SES and achievement for each half thus defined.

Other analyses that appear in some of the state-level reports investigate the differences between one group or another on various measures, as necessary to help interpret results of the

---

[9]Collinearity results from strong correlations among independent variables (i.e., size, SES, and interaction terms in this study). Since the interaction term is the product of size and SES, collinearity is a likely event (i.e., either of the factors would be likely to correlate rather highly with the product of both of them). The problem is that collinearity would inflate the error of our measures of the strength of influence (i.e., error associated with the regression coefficients). Centering drastically reduces collinearity; and, in these analyses, has successfully eliminated the problem.

regression equations. In some cases, as well, we introduce control variables (e.g., pupil-teacher ratio) to see whether such additional variables alter the prediction given by our basic model.

Montana Data

As in all these state analyses, our "sample" was planned to be all the schools and districts in each state. When an entire group, instead of a sample, is used in analyses, the calculation of significance levels is sometimes considered superfluous. The reason is that, since sampling error is not at all an issue (all cases are used, so that estimates for a subgroup are not generalized to the entire group), the observed measurements directly and accurately characterize the prevailing relationships. We have, however, retained the use of significance levels, as we believe that nonsignificant ($p>.05$) relationships, almost by definition, are practically insignificant as well. Therefore, with a few exceptions in this report, we do not provide statistics for equations with nonsignificant relationships for our size variables[10]. Understand, also, that the process of obtaining, cleaning, merging, and analyzing the data inevitably reduces the actual number of cases available for analysis by a small proportion[11]. The number of schools and districts on which we are able to base our results is somewhat less than the total number of districts in the state, but is much larger than a representative sample would be.

We contacted the Montana Office of Public Instruction (OPI) in August 1998 to obtain test score data; availability of such data for districts and schools was reported by the North Central Regional Educational Laboratory (1996). Data arrived shortly thereafter. The data set

---

[10]SES may be assumed to be negatively related to achievement in all these analyses; however, cf. the correlations in Table 14.

[11]Missing values on some data and listwise deletion of missing cases (deletion of cases that do not contain values for *all variables* in the analysis) means that the number of cases will also vary from analysis to analysis.

we received also included information about enrollment, grade configuration, and participation in school meals programs. Thus, data for all our focal variables (achievement, size, SES, and the interaction term) were obtained from the OPI.

Dependent variables (i.e., achievement test scores). All Montana students are tested at grades 4, 8, and 11 using familiar norm-referenced tests with established technical worth (e.g., Comprehensive Test of Basic Skills (CTBS); Iowa Test of Basic Skills (ITBS); and the Stanford Achievement Test). Although the Montana State Board of Education requires that local education agencies administer achievement testing, it does not require all jurisdictions to administer the *same* test, nor does it use tests results in an accountability structure (which explains *Education Week's* low accountability grade of "D," as reported in Table 2). Just three tests accounted for 86% of all case (schools): CTBS (34.2%), ITBS (31%) and Stanford (21.1%)[12]. The OPI, however, computes aggregate results by school for all tests as Normal Curve Equivalents (NCE units)[13].

The OPI does not usually release test results for its smallest schools, because the instability of such scores (due to small numbers of students taking the test) from year to year makes them unreliable measures of quality for a particular school in a particular year (Dori Nielson, personal communication, August 1, 1998). For The Matthew Project, however, the OPI released the scores on the understanding that (a) statistical analysis distributes this instability

---

[12]Differences in test construction remain problematic, however, and in addition to our major analyses using all available tests, we computed school-level regression equations for the three major tests (CTBS, ITBS, Stanford) separately as well. This procedure allowed us to see if separate analyses by test would yield similar or different results.

[13]NCE units are *equal-interval scores* with a mean of 50 and a standard deviation of about 21 (with students as the unit of analysis; variability is somewhat reduced when scores are aggregated). Use of equal-interval scores is required for the sorts of analyses conducted in the Matthew Project studies.

over all such small schools and (b) the project was not assessing nor would it report hypothetical measures of the supposed quality of any individual schools.

The OPI does not aggregate scores by district. Therefore, we aggregated tests scores to the district level ourselves. We computed the district scores as student-weighted school averages. That is, the aggregate district score was not the average of all school-level scores, but was instead the average of school-level scores weighted by the number of students in a school. This reflects an underlying principle that the aggregate district score should reflect the average of students in the district rather than the average of schools in the district.

For the Montana analyses, at both the school and district level, the we computed a composite achievement measure, the average of five achievement scores in NCE units: reading, mathematics, language arts, science, and social studies. This score, at each grade level, for both districts and schools, was our dependent variable. In addition, we computed a single measure of overall district achievement as the average school performance in the district, across all applicable grade levels, weighted by the number of students.

Independent variables (i.e., size and SES). For the Montana analyses, as in our previous work (Howley, 1995, 1996), we chose the ratio of *total school enrollment to number of grade levels* (enrollment/grade span) as our measure of school size. This measure has the advantage of controlling for the possible confounding effects of school grade span. A school enrolling 300 students in grades K-8 is arguably a smaller school than one enrolling 300 students in grades K-4.

Which measure of size to employ seemed a bit more problematic for the district level analyses, since, unlike the other states considered by the Matthew Project, Montana maintains three sorts of districts: elementary, secondary, and K-12. We resolved the dilemma by

computing the ratio of district enrollment and grade span for analyses by applicable grade level (4 and 8, 11 only, and 4, 8, and 11 for K-12 districts) and total enrollment for the composite district achievement measure.

Because our size measures, as is typical, were all positively skewed (meaning that the size distribution showed many smaller schools and fewer larger schools), we computed the natural log of all size variables in order to produce the (unskewed) normal distributions required by regression analysis.

We explored a number of possible variables for use as our SES proxy, including 1990 district-level data appearing for all US districts in the *School District Data Book* (NCES, 1995), and various more recent county-level statistics from the Bureau of the Census. None correlated with achievement so well as free-and-reduced price meals (transformed as its square root to reduce skewness). At the district level, the natural log of the percentage of American Indian students initially appeared to be a possible proxy for SES, but, in fact, with the removal of a few ($n<20$) cases (those with the highest proportion of Indian students), the variable proved to be irremediably skewed, and was abandoned as a suitable SES proxy for regression analyses.

Unfortunately for this research, many of Montana's smallest schools do not provide meals programs (they are too small to support the employment of kitchen staff). The practical upshot of this circumstance is that there is no free-and-reduced-price-meals rate available for these

schools. Of the 234 smallest elementary schools, 112 do not offer meals programs. The average total school enrollment was about 36, for the most common K-8 configuration among these schools. The 122 smallest schools *with* meals programs enrolled, on average, about 90 students in grades K-8[14].

Results

We report results in two main sections: (1) regression results and (2) results of correlations between SES and achievement by size halves (the unit divided at the median size). Regression results tell us something about excellence (i.e., the size conditions needed to maximize student achievement), whereas the correlations by size halves tell us something about equity (the strength of the bond between achievement and SES as size conditions vary). Within each section, results are reported first for school-level analyses and second for district-level analyses. The presentation of results concludes with a summary of ancillary analyses: regression analyses for districts enrolling at least 25% American Indian students, critical correlations with a range of alternative SES proxies, and regression analyses that add pupil- teacher ratio to the basic model--in order to investigate the possible confounding effects of class size differences between schools.

Regressions (schools). We regressed each of the three levels of achievement scores (grade 4, grade 8, and grade 11) on the pertinent size, SES, and interaction terms (i.e., centered

---

[14]Mean 4th and 8th grade NCE scores, respectively were 58.2 and 61.7 (*without* meals rates) and 57.2 and 57.3 (*with* meals rates). In one-way analyses of variance these differences were not statistically significant (p=.685 and p=.074, respectively), but because we are dealing with the universe of such schools, it would be wise to conclude that a practical difference equivalent to +0.4 standard deviations in test scores exists among 8th grade students in the smallest schools (5E and 6E OPI size categories) without meals programs as compared to such schools with meals programs.

logged cohort size, centered square-rooted free-and-reduced-price-meals rate, and the product of those two centered variables). Table 5 presents the school-level results.

Table 5
Summary of Hierarchical Regression Analysis (Grades 4, 8, and 11)

Grade Level	Variables in the Equation	B	SE B	β	p	adj R^2	N[a]
4	SIZE	+ .588	.438	.067	.181	.249	325
	SES	-2.460	.449	-.526	.000		
	Interaction	- .583	.167	-.183	.001		
8	SIZE	+ .061	.442	.009	.891	.182	220
	SES	-2.274	.326	-.484	.000		
	Interaction	- .380	.205	-.128	.065		
11	SIZE	-1.123	.612	-.161	.068	.297	168
	SES	-2.684	.326	-.587	.000		
	Interaction	- .034	.308	-.010	.913		

Notes. a. listwise N

The results given in Table 5 show a weak small positive *direct* effect of size for grade 4 (+.588) and grade 8 (+.061) . If we were using a sample instead of the universe of schools we would conclude that these observed effects were not statistically significant. At grade 11, however, the direct effect of size is somewhat stronger and it is negative rather than positive; statistical significance, however, would also be judged non-significant (p=.068).

The *indirect* effects of size, as reflected in the interaction term, are all in the hypothesized direction (i.e., negative). The grade 4 interaction effect is sufficiently strong that it would be judged as highly statistically significant (p=.001) under the sampling framework. The grade 8 effect is weaker (p=.065), and the grade 11 effect, though negative, would be without influence in practical terms, and is statistically quite nonsignificant (p=.891).

Finally, with each of these equations, the independent variables account for less than 30

BEST COPY AVAILABLE

percent of the variance in achievement (similar to our findings reported for Texas and West Virginia, but lower than reported for California, Ohio, and Georgia). Still, one equation in Table 5 (for grade 4) does indicate the presence of a significant interaction influence.

As noted previously, Montana does not require all schools to administer the same test. Therefore, we repeated the full-sample analyses for subgroups of schools at the various grade levels according to the test used (i.e., CTBS, ITBS, or Stanford). Table 6 shows the results for any equation in which we observed either a substantial significant direct or indirect (interactive) influence of size. The interaction effects turn out to be somewhat more marked than in the analyses combining all tests.

Table 6
Summary of Hierarchical Regression Analysis(by test)

Grade Level	Test	Variables in the Equation	B	SE B	β	p	adj R^2	N[a]
4	ctbs	SIZE	-1.950	.719	-.203	.008	.330	121
		SES	-2.067	.366	-.423	.000		
		Interaction	-1.580	.334	-.354	.000		
4	itbs	SIZE	+1.265	.569	+.160	.028	.350	128
		SES	-2.572	.316	-.639	.000		
		Interaction	-1.087	.215	-.395	.000		
4	stanford	SIZE	-1.043	1.144	- .130	.365	.315	70
		SES	-2.385	.565	- .487	.000		
		Interaction	+ .759	.376	+ .283	.048		
8	itbs	SIZE	- .796	.539	-.153	.144	.183	78
		SES	-2.007	.475	-.526	.000		
		Interaction	- .845	.280	-.376	.004		
11	stanford	SIZE	-5.374	1.675	-.766	.003	.187	34
		SES	- .986	.827	-.224	.242		
		Interaction	-1.467	.728	-.484	.053		

Notes. a. listwise N

BEST COPY AVAILABLE

The data reported in Table 6 are generally consistent with the analyses irrespective of test, with several interesting differences. First, statistically significant interaction effects appear in all three tests at the fourth grade level. Interestingly, the sign is reversed in the case of the 70 schools using the Stanford Achievement Test to assess grade 4 achievement[15]. Second, the direct effect of size reverses sign for the 128 schools using the Iowa Test of Basic Skills, and it is negative but nonsignificant for the Stanford schools. Third, significant interaction effects are also evident for the 78 schools using the ITBS for grade 8 and the 34 schools using the Stanford for grade 11. Finally, above grade 4, our model accounts for rather less variance (i.e., about 18 percent of the variance in achievement) than it does at grade 4.

Because Montana, like other states, exhibits preference for one or another school configuration types, we applied our regression model to subgroups of the most popular configurations at the elementary, middle, and high school levels. We had complete data for 176 K-6 elementary schools, 172 7-8 schools, and 167 9-12 high schools. Performing the analyses in this way imposes an additional control related to the variability in grade configuration. In none of these regression equations did the interaction term prove significant, and only at the grade 8 level did the observed negative coefficient of size prove statistically significant (β= -.146, p=.030).

---

[15]These schools are smaller (24 vs 32 grade 4 cohort size) and enroll proportionally fewer (about 5% versus about 10%) Indian students than schools that give the CTBS and ITBS. Meals rates are approximately the same, as is locale (rural vs. urban). The interaction term in this equation, unlike the others, is quite sharply skewed, rendering the observed results dubious (marked skewness violates a fundamental assumption of regression analysis, namely that variables be normally distributed). For this reason, Table 7 does not report ESs for this equation.

Table 7 reports effect sizes (ES) [16] associated with changes in the underlying SES measure (free-and-reduced-price-meal rates), also expressed as a percentile[17]. A positive ES means that the influence of size on achievement at that SES level tends to increase aggregate school-level achievement, whereas a negative ES means that the influence of school size at that SES level tends to decrease aggregate school-level achievement. The larger the number, the stronger the effect, whether negative or positive. The ES's in the various school-level tables (including Table 7) show the change in achievement associated with a change of about 16 students from the average grade 4 cohort size of about 24 students.

Table 7
Effect Sizes of School Enrollment on NCE Mean Scores (Grade 4)
for Varying Levels of School SES

SCHOOL SES		EFFECT SIZES BY MODEL		
meals rate percentiles	meal rates	grade 4 (all tests)	grade 4 (CTBS only)	grade 4 (ITBS only)
1.00	.00	+.39	+.75	+.96
5.00	7.15	+.21	+.31	+.59
10.00	15.86	+.12	+.09	+.41
15.00	19.42	+.09	+.02	+.35
20.00	23.29	+.06	-.05	+.29
30.00	28.86	+.03	-.14	+.22
40.00	33.33	+.00	-.20	+.16
50.00	38.18	-.03	-.27	+.11
60.00	42.31	-.05	-.32	+.06
70.00	48.33	-.08	-.40	+.00
80.00	55.72	-.11	-.48	-.07
85.00	61.61	-.14	-.54	-.12
90.00	75.70	-.19	-.68	-.24
95.00	94.83	-.26	-.85	-.38
99.00	100.00	-.28	-.90	-.42

[16]Effect sizes of .20-.30 are considered moderate and practically significant. Effect sizes of .50 and higher are considered quite strong. As a very rough practical guide, one might regard an ES of +1.0 as a comparative improvement of about one year of learning and an ES of -1.0 as a comparative loss of about one year of learning.

[17]Knowing the percentile rank of our SES measure allows readers to judge the extent of the grade 4 effects among Montana schools. For instance, in Table 7, 15 percent of schools have meals rates of 61.61or higher.

BEST COPY AVAILABLE

Together, the findings suggest that, overall, the interaction effect, though observable, is *not* strongly influential among Montana schools. Only at the fourth grade level does the interaction effect contribute to explaining what is, in fact, a comparatively modest amount of variance in student achievement (i.e., 18.2% to 35%, see Tables 5 and 6). Montana schools serving impoverished communities exert a somewhat more positive influence on achievement on grade 4 achievement if they are smaller; and Montana schools serving distinctly affluent communities exert a somewhat more positive influence on grade 4 achievement if they are larger. This influence is not strongly evident at grade 8 and 11, however. Finally, it is not evident at *any* level when schools are analyzed by the most popular grade configurations in Montana.

Regressions (districts). Montana maintains elementary, secondary, and K-12 district configurations. We computed equations for each grade level and for districts as a whole (averages of schools within districts, weighted for enrollment at the school-level). We also computed the equations with cases segregated by district type, to control for the differences that might be associated with organizational differences.

As readers might infer from Table 1, there is something peculiar about many Montana districts--a situation that would surprise educators as well as citizens in many other parts of the nation. Many of Montana's districts operate a single school, and usually, such schools are themselves small or very small by national standards. Of 202 elementary districts in our data set, 53 operate a single school; of 112 high school districts in our data set, 107 are one-school districts[18].

We computed (1) four regression equations for all districts together (whole district

---

[18] That is, just 5 high school districts operate more than 1 HS; whereas 149 elementaries operate more than 1 school; interestingly, just 29 elementary districts represented in our data set operate more than 2 schools.

composite achievement, grade 4, grade 8, and grade 11), as well as (2) three regression equations for elementary districts (district average, grade 4, and grade 8) (3) one regression equation for high school districts (grade 11), and (4) four regression equations for K-12 districts (whole district, grade 4, grade 8, and grade 11). Of these 12 equations, the interaction term was significant in just two equations; results are summarized in Table 8 and the related effect sizes (paralleling those in Table 7) are given in Table 9.

Table 8
Summary of Hierarchical Regression Analysis (Districts)

Grade Level	Variables in the Equation	B	SE B	β	p	adj R^2	N[a]
8 (elementary districts)	SIZE	-1.204	.498	-.181	.017	.292	147
	SES	-2.290	.323	-.506	.000		
	interaction	- .794	.278	-.209	.005		
whole district (K-12 districts)	SIZE	-1.737	.844	-.260	.045	.261	50
	SES	-1.356	.648	-.307	.042		
	Interaction	-1.657	.801	-.310	.044		

Notes. a. listwise N

BEST COPY AVAILABLE

Table 9
Effect Sizes of Enrollment on NCE Mean Scores (Districts) for Varying Levels of District SES

DISTRICT SES		EFFECT SIZES BY MODEL	
meals rate percentiles	meal rates	grade 8 (elementary)	whole district average (K-12 district)
5.00	9.40	+.14	+.40
10.00	14.30	+.05	+.23
15.00	15.80	+.03	+.18
20.00	18.50	-.01	+.09
25.00	20.20	-.03	+.05
30.00	22.20	-.06	-.01
35.00	24.60	-.09	-.07
40.00	26.00	-.11	-.10
45.00	28.30	-.13	-.16
50.00	30.60	-.16	-.21
55.00	33.00	-.17	-.24
60.00	34.10	-.19	-.29
65.00	38.70	-.24	-.38
70.00	41.70	-.27	-.44
75.00	46.00	-.31	-.52
80.00	51.00	-.35	-.61
85.00	58.40	-.41	-.73
90.00	69.50	-.49	-.91
95.00	91.20	-.64	-1.21
100.00	100.00	-.69	-1.32

Note. Grade 8 ESs show the proportional change in achievement for every change of about 15 students from an average grade 8 district cohort size of about 19 students (anti-log bases); the equivalent statistics for the K-12 model are a change of about 125 students from an average district enrollment of 227 (anti-log bases).

Critical correlations (schools). The school-level regression equations about which we provide information in Tables 5 -7 relate school size to *levels* of student achievement. Another view of the patterns hypothetically evident in these data, however, concerns the *equity* of achievement. What might that be? Most people understand inequity in school finance. Some schools receive dramatically less funding than others; frequently schools serving impoverished communities are badly funded and affluent communities are well funded (Kozol, 1991). In simple terms, the source of inequity is economic power, and something akin to SES (call it wealth, given the reliance on property tax revenues to fund schooling) determines the level of school funding. Improvements in financial equity would require better funding of schools

BEST COPY AVAILABLE

serving impoverished communities, that is, such improvements would require that we *break* the link between wealth (or SES) and school finance. Why *should* the rich enjoy the best schools?

Equity in achievement represents the operation of the same law. Which children, in general, enjoy the highest achievement? More affluent children do. Some observers, of course, believe that since affluence and ability correlate well, this state of affairs is actually very fair. Others, however, noting that among the affluent and the impoverished a great range of abilities exist, and that in any walk of life a similarly great range of abilities persists, believe that the low achievement of impoverished children is not nearly so fair as it at first might seem (e.g., Howley, Howley, & Pendarvis, 1995). In this view, public schooling can and should do much more to nurture the learning of all students. As with financial equity, equity in achievement means breaking--or at least substantially mitigating--the prevailing bond between SES and achievement[19].

The existence (even the weak existence) of an interaction effect in Montana suggests, together with the small scale of the Montana state educational system, suggests the possibility of an equity effect associated with size. One way to test the existence of such an effect is to divide Montana schools into two groups, those below the median size and those above the median size and to calculate, within each group, correlations (Pearson r's) between SES and aggregate school achievement. The square of the correlation (R^2) will give the proportion of variance in achievement accounted for by SES. Comparing the values of R^2 in the smaller versus the larger groups of schools will give us an idea of the degree of mitigation, if any, exhibited by the small

---

[19]In practical terms, one is unlikely to break the bond completely, because the negative effects of poverty can be eliminated only when a society finds them intolerable and actively cultivates the well-being of the poor. Even in the current economic boom, however, such a realization has not overtaken the US, and in general, the gap between the affluent and the impoverished is growing ever wider here. Also, some observers balk when they realize that breaking the bond must apply not just to the poor, *but to the affluent as well.*

schools. Tables 10-12 report information about these relationships in smaller versus larger schools in Montana.

Table 10
Grade 4 Achievement, Socioeconomic Status, and School Size in Two Groups of Montana Schools

	means and correlations				
Size	achievement mean[a]	meal rate mean	grade 4 enrollment	$r_{y,z}$	$r^2_{y,z}$
larger[b] schools	56.5	39.1%	52.0	-.67**	.45
smaller[c] schools	57.1	44.5%	8.5	-.33**	.11

Notes. $R^2_{y,z}$ is the variance in y (SES) associated with z (achievement). $R_{y,z}$ is the correlation.

** => p < .01 (two-tailed)

[a] all tests, all cases (including those smaller school with missing data on SES variable, see note c; achievement mean *excluding* those cases = 56.5)

[b] n = 195

[c] n = 130 (66 cases with missing values on SES variable)

Table 11
Grade 8 Achievement, Socioeconomic Status, and School Size in Two Groups of Montana Schools

	means and correlations				
Size	achievement mean[a]	meal rate mean	grade 8 enrollment	$r_{y,z}$	$r^2_{y,z}$
larger[b] schools	57.3	33.8%	88.5	-.66**	.45
smaller[c] schools	58.2	44.6%	7.9	-.31**	.10

Notes. $R^2_{y,z}$ is the variance in y (SES) associated with z (achievement). $R_{y,z}$ is the correlation.

** => p < .01 (two-tailed)

[a] all tests, all cases (including those smaller school with missing data on SES variable, see note c; achievement mean *excluding* those cases = 56.6)
[b] n = 129
[c] n = 91 (44 cases with missing values on SES variable)

Table 12
Grade 11 Achievement, Socioeconomic Status, and School Size in Two Groups of Montana Schools

	means and correlations				
Size	achievement mean[a]	meal rate mean	grade 11 enrollment	$r_{y,z}$	$r^2_{y,z}$
larger[b] schools	56.2	23.1%	127.2	-.52**	.27
smaller[c] schools	55.7	34.8%	15.2	-.58**	.34

Notes. $R^2_{y,z}$ is the variance in y (SES) associated with z (achievement). $R_{y,z}$ is the correlation.

** => p < .01 (two-tailed)
*** => p < .001 (two-tailed)

[a] all tests

[b] n = 83
[c] n = 85

Tables 10-12 suggest that an equity effect of size exists in Montana schools for grades 4 and 8, but not for grade 11. For grade 11, the association is approximately the same in larger as compared to smaller schools. As footnote 14 suggests, Montana's smallest schools (OPI size categories 5E and 6E) that do not have school meals programs exhibit comparatively higher achievement than the smallest schools with lunch programs, and this fact is reflected in the achievement statistics reported for the smaller grade 4 and 8 schools in Tables 10 and 11. Because the meals rate data of course do not include these schools, we have also reported (note c in Tables 10 and 11) the achievement mean excluding these schools. In any case, despite small differences in achievement one way or another, the key point to take from Tables 10 and 11 is that achievement is at least comparable, despite the lower SES in the smaller schools. The means data amplify the implication from correlational analysis: these small schools appear to be mitigating the negative influence of poverty on achievement.

Critical correlations (districts). We repeated the correlations by size halves for districts. We analyzed the data for all districts together and segregated by district type. Specifically, our analyses compared smaller and larger halves (1) for all districts regardless of configuration (i.e., configuration as elementary, high school, or K-12 districts) [a] overall achievement and [b] grade 4 achievement; (2) for elementary districts only [c] overall achievement, [d] grade 4 achievement, and [e] grade 8 achievement; (3) for high school districts only [f] grade 11 achievement; and (4) for K-12 districts only [g] overall achievement, [h] grade 4 achievement, [i] grade 8 achievement, and [j] grade 11 achievement. We also performed the analysis for [k] all N=210 districts with grade 8 students and with complete SES, size, and achievement data. For this final analysis we simply selected the cases meeting that criteria and divided the group in half

according to the median for that group on size (24.5). Grade 8, moreover, is the last year in the Montana data set before the dropout year (grade 10). Because disaffected students leave by the next testing year (grade 11), grade 8 is the year in which any equity disparities by size are likely to be sharpest.

This series of analyses provides 11 comparisons of the association of SES and achievement in Montana districts. In 9 cases, the observed differences favored smaller districts. The sharpest differences are found in the two grade 8 analyses[20], but also in all four analyses of the relationship in K-12 districts (analyses g-j). The K-12 findings are all the more remarkable because sample size is comparatively small (i.e., comparisons groups of just 25 cases). Tables 13 and 14 report the two grade 8 and the four K-12 district findings.

---

[20]That is (1) elementary districts only, and (2) all districts with grade 8--analyses "e" and "k."

Table 13
Grade 8 Achievement, Socioeconomic Status, and District Size in Four Groups of Montana Districts

		means and correlations				
	group	achievement means[a]	meal rate mean	mean district enrollment	$r_{y,z}$	$r^2_{y,z}$
Size						
larger districts	elementary only[b]	57.1	36.3%	838.3	-.60***	.36
	all districts[c]	56.4	35.5%	924.9	-.64***	.41
smaller districts	elementary only[b]	60.1	47.0%	57.2	-.35***	.12
	all districts[c]	59.4	41.3%	98.6	-.39***	.15

Notes. $R^2_{y,z}$ is the variance in y (SES) associated with z (achievement). $R_{y,z}$ is the correlation.

*** => p < .001 (two-tailed)

[a] all tests
[b] n = 98 (larger group) and n = 50 (smaller group; 48 cases with missing data on meals rate)
[c] n = 104 (larger group) and n=106 (smaller group); all district group with complete grade 8 data divided at group median on grade 8 cohort size [24.5])

BEST COPY AVAILABLE

Table 14
Achievement, Socioeconomic Status, and District Size in Two Groups of Montana K-12 Districts

Size	achievement measure	achievement means[a]	meal rate mean	mean district enrollment	$r_{y,z}$	$r^2_{y,z}$
		means and correlations				
larger K-12 districts[b]	grade 4	56.0	32.9%	586.2	-.57**	.32
	grade 8	57.2			-.56**	.31
	grade 11	54.6			-.66**	.43
	district	55.9			-.67**	.45
smaller K-12 districts[c]	grade 4	56.3	32.8%	130.1	+.07(ns)	.00
	grade 8	58.6			-.34(ns)	.12
	grade 11	57.8			-.17(ns)	.03
	district	57.5			-.22(ns)	.05

Notes. $R^2_{y,z}$ is the variance in y (SES) associated with z (achievement). $R_{y,z}$ is the correlation.

** => p < .01 (two-tailed)
ns => not significant (two-tailed)

[a] all tests
[b] n = 24 (grade 4); 25 (grade 8); 24 (grade 11); 25 (district)
[c] n = 25 (grade 4); 26 (grade 8); 26 (grade 11); 26 (district)

Table 13 uses two groups of districts with grade 8 achievement data. The elementary group, as in other analyses, excludes a substantial number of cases due to missing meals rate (SES) data. The group of districts with complete data, divided at the median of that group on size, however, gives nearly identical results. Table 14 shows the strongest differences in these series of analyses, a result that is all the more remarkable given the quite small number (n=51) of K-12 districts in Montana. By national standards, these are not large districts; the smallest of the

BEST COPY AVAILABLE

"smaller" group of K-12 districts has a total K-12 membership of 49 students, but the largest of the large group has a total K-12 membership of just 2,078 students. Though 2,000 students would be considered a rather large district in parts of the western US, in other regions of the nation (e.g., regions where whole-county districts are common), 2,000 would be considered *small*. The relationship between SES and achievement is so reduced as to be negligible in three cases out of four, with SES still accounting for a mere 12% of the variance in student achievement at the grade 8 level. And for weighted district achievement as a whole, the relationship in small districts accounts for a non-significant 5% of variance in achievement, whereas it accounts for a quite significant ($p<.01$) 45% in larger districts.

Additional analyses. We performed several ancillary analyses, which we will summarize here briefly. First, no evidence of any size effect (direct or indirect) was found among the 47 Montana districts with a high proportion (greater than or equal to 25% of enrollment) of American Indians; in fact *none* of the variables in the equation (including SES) was a significant predictor of overall district-level achievement. Indians in Montana attend schools and districts about the same size as other students. This eventuality may be partly due to the fact that Indians are the nation's most rural population group, with about 50% of Indians nationally living in rural areas, but the Montana treatment of minorities in this regard nonetheless contrasts markedly with the situation of minorities in other states, where concentrations of African Americans and Hispanics are apt to attend the *largest* schools in the *largest* districts.

Second, using a range of SES proxies[21] we calculated the relationship between both school-level achievement and SES and district-level achievement and SES by size halves. Even

[21]These include such district- and county-level measures as were available: median household income, unemployment rate, percent minority children, percent children living in poverty, percent of the over-25 population with fewer than 12 years' schooling, and so forth.

though these measures had proven to be less workable for the main analyses as SES proxies than meal rate data, this additional series of correlational analyses clearly showed an equity effect at both the school and district level completely consistent with the reports found in Tables 10-14. Explained variance in the smaller schools and districts was nearly always a fraction of the variance explained in larger schools and districts (often 1/10-1/3 the variance).

Finally, we investigated the possibility that the observed interaction effects might be attributable to the small class sizes associated with smaller schools. To test this possibility we recomputed the school-level regression equation terms (see Tables 5 and 6) for which we had found significant interaction terms[22], this time with the pupil-teacher ratio entered in the equation first as a control for class size. SPSS output for these equations appears in Appendix B. In each case, the regression coefficient of this new variable was nonsignificant. *We conclude that those instances of size effects which we can report for Montana are distinct from any that might be associated with variability in class size.*

Limitations

All studies have limitations. For the Montana study, several limitations have already been noted in passing. First, Montana schools administer various and different tests at the three grade levels at which the Board of Education mandates testing. Differences between the tests basically suggest an unknown degree of incomparability, doubtless resulting in some error. To compensate for this threat to validity, we calculated school-level regression equations by test. These analyses yielded results fundamentally consistent with those for all tests combined, with some variation in regression coefficients, but not in explained variance.

---

[22]grade 4 (all tests), grade 4 (CTBS, ITBS, Stanford), and grade 8 (ITBS)

Second, only a small number of Montana districts encompass all grades K-12. Most districts are separately configured elementary and high school districts. To compensate for this distinction, we calculated our district-level regression equations by district type, as well as for all districts combined.

Third, our SES measure is known for its weakness in comparison to carefully constructed measures. Rates of free-or-reduced-price meals are subject to variation unrelated to SES, such as the tendency of rural people to access social services less frequently than urban people, the vigor with which school staff solicit applications, and the tendency of high school students not to participate in the meals program. Further, in Montana in particular, many of the smallest schools do not maintain meals programs. The likely effect of this latter threat is a reduction in the direct and indirect effects of size at the grade 4 and grade 8 levels; the equations in Tables 5 and 6 may underestimate the differential influence of size at these grade levels. And the undersubscription to the meals program may, for grade 11, produce a similar effect. This shortcoming is a common one in studies such as this; meals rate data are commonly the only school-level proxy for SES available to researchers.

Conclusions and Discussion

The primary aim of this study was to investigate the hypothetical interactive effects of school and district size on aggregate student achievement in Montana--a sort of differential "excellence effect" of size. A secondary aim concerned testing the hypothetical equity effect of size in disrupting the usual strong negative relationship between SES (meals rate) and achievement.

Interactive effects. We found limited evidence in Montana of the existence of the

interaction effect of school and district size (see Tables 5-9). We found some evidence of an interaction effect among K-12 districts, but only for our summative measure of whole-district achievement (i.e., weighted means of all composite grade-level measures), and we found a district-level interaction effect for grade 8, but only among elementary districts.

The evidence for the existence of an interaction effect seems strongest at the school level for grade 4, where a reliable interaction effect was statistically significant in three equations (see Tables 5 and 6). The amount of variance explained by our regression equations in Montana, however, was rather modest, varying between 18% and 35%. In similar achievement studies explained variance ranges up to 60% or higher.

At the school level, evidence of an interaction effect of size appears somewhat stronger when achievement scores are analyzed by tests rather than with all tests combined. Anomalies by different test, however, show up in Table 6. In particular, it seems that questionable results from the minority of schools (n=70) administering the Stanford Achievement Test may somewhat obscure the strength of the grade 4 interaction among the 249 schools administering the CTBS and ITBS to their students. In those 249 schools, the interaction effect is approximately 3 or 4 times as strong as reported for the analysis using all tests combined (unstandardized regression coefficients of -1.580 and -1.087 in Table 6 compared to -.583 in Table 5). Whatever the anomalies, however, *we conclude that in Montana, even the strongest evidence of an interaction effect is comparatively weak.*

What, aside from the limitations previously delineated, might account for the weak evidence of an interaction effect in Montana[23]? The most likely answer, from a statistical

---

[23]Evidence from California (Firedkin & Necochea, 1988), West Virginia (Howley, 1995, 1996), Georgia (Bickel, 1999a), Ohio, (Howley, 1999), and Texas (Bickel, 199b) show considerably stronger evidence of an interaction effect across units of analysis (schools and districts) and grade levels.

perspective, is reduced variance of size in Montana. From a practical and policy perspective, this means simply that--as is evident in Table 1--Montana maintains small school and small districts, in comparison to national norms. The average K-6 elementary school in Montana enrolls 165 students, and the median (the size in the exact middle of the K-6 group) is just 96. The average 7-8 school (the norm according to our data set for middle-level schools in Montana) enrolls 165 students, with a median of 122. The real surprise is the high schools, however, with an average total enrollment of only 275 students and a median of just 114 students--total-- in grades 9-12. In fact, 75% of Montana high schools enroll fewer than 300 students, and that figure is about the same for K-6 schools; that is, 75% of Montana elementary schools *also* enroll fewer than 300 students. In our data set, the smallest Montana school (in 1996) enrolled a single student in the 8th grade, and the largest enrolled 2043 students (an urban high school). Large high schools are unusual in Montana, of course, with just 7% enrolling more than 840 students in grades 9-12.

Given this reduced variability it is rather surprising to find any indication of an interaction effect of school and district size in Montana. The fact that such an effect is nonetheless weakly evident may tell us more about the prevalence of the phenomenon of the interaction than it does about Montana. Its discovery in Montana might suggest that the interaction of size and SES is a fundamental characteristic of the structure of systems of schooling, at least in North America, where political-economic choice has fostered the creation of larger units over the course of the 20th century.

Equity effects. Though identifiable interaction effects were found in Montana, the much more impressive finding is the consistent presence of an equity effect of small school and district size. At all grade levels, and at both the district and school levels, the association between SES and aggregate student achievement is weakened in smaller units. At the district level, in fact, the

results are dramatic for grade 8 (the testing cycle before the dropout year) and for all grade levels (including the whole-district composite score) in K-12 districts.

Even in Montana, where small schools and districts prevail, the evidence in favor of an equity effect of small-scale schooling is strong. Perhaps the most remarkable finding to come from this study is that achievement in poorer, smaller schools and districts equals, *if not exceeds*, achievement in markedly more affluent, larger schools and districts. The strongest evidence that "smaller is much better" for students from impoverished communities appears in Table 14, where the n=210 districts with grade 8 scores are examined, based on district size divided at the median. Here we see a difference of about 3 NCE units, which is an effect size of about 1/3 to ½ standard deviation units[24], in favor of smaller, poorer districts, and a sharply reduced association of SES and achievement in smaller versus larger districts.

In a sense, the "real test" of both the excellence hypothesis (interaction effect) and the equity hypothesis in Montana at the district level accrues to the K-12 districts. The reason for this claim is that the "normal" or "most common" view of a school district is that it encompasses a system that serves all grades K through 12. As we have seen, Montana frustrates the common expectation by being composed principally of elementary and secondary districts. Here, overall, we see evidence that would confirm both effects: for composite district achievement overall (as in Table 9) and in larger versus smaller K-12 districts (in Table 13). The argument that this analysis is the "best case" analysis is somewhat strengthened by the fact that K-12 districts are as likely to be found in rural areas and small towns as elementary and secondary districts; the K-12 district is not an urban phenomenon in Montana. Here we see a comparable degree of affluence in both groups of districts (meal rate of about 33%, a bit higher than the median for districts of

[24]Each level of aggregation loses variance, and at the district level, the standard deviation is about 1/3 the individual-level standard deviation of 21 (typically 7-9 NCE units in these data

30.5%), and slightly higher achievement in the smaller districts, but the surprising finding is that, across all four measures of achievement (grade 4, grade 8, grade 11, and whole-district composite), the association of SES and achievement in the larger districts accounts for about 40% of the variance in achievement, but almost nothing (all correlations are nonsignificant) among the smaller districts.

Final observations. The evidence suggests that Montana has done what educational researchers, the present team included, believe to be impossible: accumulated data that, when analyzed in the conceptual framework of the Matthew Project, provide evidence that groups of more impoverished students have equaled, if not out-performed, groups of more affluent students *on standardized, norm-referenced measures of achievement*. Except for the ancillary analyses, discussed in the "Results" section, all analyses have included the comparatively lower-performing districts and schools serving large proportions of American Indian students (about whose educational fate Montanans ought to worry). *These results are not, then, the result of excluding Montana's least privileged students and communities from the analysis. This fact makes the results are the more remarkable.*

Though causal inferences in statistical analysis are dubious ventures, politicians and some SEA officials do imagine that state-level accountability schemes (as adopted recently in many states) will cause school improvement, particularly causing rises in levels of student achievement to meet "high standards." Montana gets a D from *Education Week* for not having such a scheme in place. Yet it leads the nation in NAEP results, it maintains the small schools and small districts thought elsewhere to be effective but prohibitively expensive, and it seems to enjoy arrangements that particularly cultivate higher levels of achievement in its poorest communities. We would be hard pressed to *prove* that the commitment to small-scale schooling is responsible

for these achievements, including the "D" for poor accountability, but the evidence probably points to relationships that must be hypothesized as considerably more than coincidental. At the very least, Montanans would be well-advised to retain their commitment to small schools and districts.

If that commitment entails a low level of tolerance for state accountability schemes, then Montana may have to accept continued poor grades for accountability as the price of a modicum of excellence and equity. It *would* seem unwise to trade the latter for the former.

References

Bickel, R. (1999a). *The Matthew Project: State report for Georgia.* Randolph, VT: Rural Challenge Policy Program.

Bickel, R. (1999b). *The Matthew Project: State report for Texas.* Randolph, VT: Rural Challenge Policy Program.

Clones, D. (1998). *Development report card for the states* (12th ed.). Washington, DC: Corporation for Enterprise Development.

Cronbach, L. (1987) Statistical tests for moderator variables: Flaws in analyses recently proposed. *Psychological Bulletin, 102*(3), 414-417.

Dorn, S. (1998). The political legacy of school accountability systems. *Education Policy Analysis Archives, 6*(1) [On-line serial]. (Available World Wide Web: http://olam.ed.asu.edu/epaa/v6n1.html)

Education Week. (1998). *Technology counts: Putting school technology to the* test (*Education Week, 18*(5), special issue). Washington, DC: Author and the Milken Exchange on Education Technology.

Education Week. (1997). *Quality counts: A report card on the condition of public education in the 50 states* (supplement to *Education Week*, January 22, 1997, Vol. XIV). Washington, DC: Author and the Pew Charitable Trusts.

Friedkin, N., & Necochea, J. (1988). School system size and performance: A contingency perspective. *Educational Evaluation and Policy Analysis, 10*(3), 237-249.

Howley, C. (1999). *The Matthew Project: State report for Ohio.* Randolph, VT: Rural Challenge Policy Program.

Howley, C. (1996). Compounding disadvantage: The effects of school and district size on student achievement in West Virginia. *Journal of Research in Rural Education, 12*(1), 25-32.

Howley, C. (1995). The Matthew principle: A West Virginia replication? *Education Policy Analysis Archives* [On-line serial] *3*(18). (Available World Wide Web: http://olam.ed.asu.edu/epaa/v3n18.html)

Howley, C., Howley, A., & Pendarvis, E. (1995). *Out of our minds: Anti-intellectualism in American schooling.* New York: Teachers College Press.

Huang, G., & Howley, C. (1993). Mitigating disadvantage: Effects of small-scale schooling on student achievement in Alaska. *Journal of Research in Rural Education, 9*(3), 137-149.

Johnson, F. (1989, March). *Assigning type of locale codes to the 1987-88 CCD public school universe*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA. (ERIC Document Reproduction Service No. ED 312 113)

Kahlenberg, R. (1999, March 31). Commentary: Economic desegregation. *Education Week* p. 52-30.

Mehrens, W. (1998). Consequences of assessment: What is the evidence? *Education Policy Analysis Archives, 6*(13) [On-line serial]. (Available World Wide Web: http://olam.ed.asu.edu/epaa/v6n13.html)

National Center for Education Statistics. (1995). *School district data book* [CD-ROM]. Washington, DC: Author. (Available World Wide Web: http://nces.ed.gov/surveys/SDDB/introd.htm)

National Center for Education Statistics. (1998). *Overview of public elementary and secondary schools and districts: School year 1997-98* (Tables 1, 4, and 6). Washington, DC: Author. (Available World Wide Web, July 1999: http://nces.ed.gov/pubs99/1999322/)

North Central Regional Educational Laboratory. (1996). *State student assessment programs database*. Chicago: Author.

Schwinden, T., & Brannon, L. (1993). *School reorganization in Montana: A time for decision*. Helena, MT: Montana School Boards Association.

Weber, M. (1947). *The theory of social and economic organization (Translation of Wirtschaft und Gesellschaft*, T. Parsons, trans.). New York: The Free Press.

APPENDIX A

Note on Method

This study employs a very simple regression model, with an interesting methodological innovation. The model was developed and applied by Noah Friedkin and Juan Necochea (1988) to a California data set. The study is widely cited in the school size literature. The Friedkin and Necochea model involves regressing achievement on size, SES, and the product of SES and size (interaction term). If a significant interaction effect is found, effect sizes (ESs)at varying levels of SES can be calculated by taking the partial derivative of the regression equation. Standardizing the partial derivative and evaluating the resulting expression at chosen SES intervals gives the ES of size on achievement at those SES levels.

The partial derivative is a procedure in calculus and constitutes one of the original features of this method. Any use of calculus in educational studies is unusual, and hence merits careful explanation, which is provided in the following paragraphs.

In calculus, the derivative (often written as "dy/dx") gives the value of the change in a dependent variable, y, associated with change in a single independent variable, x. The derivative is a ratio, often construed as a generalized form of slope (familiar in algebra as the ratio of rise over run, or $[y_2 - y_1]/[x_2 - x_1]$). The regression coefficient, of course, provides an estimate of slope in regression analysis, with regression coefficients appearing as constants in regression equations (during the process of taking a derivative it is important to distinguish which values are constant and which vary).

Unlike the derivative, the partial derivative (like the partial regression coefficient) is particularly useful in working with equations with two or more independent variables (e.g., as in regression equations). The partial derivative merely gives the rate of change in the function (i.e., the value of the dependent variable) with respect to one independent variable as another is held constant. This resembles the way in which partial regression coefficients give the influence of one variable on another when the influence of a third variable is eliminated.

To calculate the partial derivative, one variable (either x or y) is held constant while differentiation proceeds with respect to the other; the variable "held constant" is treated during differentiation as if it were a constant. Afterwards, values of the variable held constant (socioeconomic status, in this case) can be substituted in the resulting equation to calculate actual values of the partial derivative function (i.e., to determine the influence of the variable not held constant on the dependent variable with respect to differing values of the variable that was held constant during partial differentiation). In this study the effects of size on achievement are hypothesized to vary by socioeconomic status, and the partial derivative hypothetically provides a mechanism to evaluate the differences.

For this study, the applicable partial derivative will give the effect of change in size (defined as cohort-level enrollment or total district enrollment as applicable) on achievement (test scores) as socioeconomic status (free-and-reduced-price-meal rates) is held constant. The general form of the mathematical model proposed is given by the following equation:

$$f(z) = ax + by + cxy$$

(equation 1)

where

a, b, and c are the unstandardized regression coefficients (of size, socioeconomic status, and the interaction term, respectively);

z represents values of achievement (the dependent variable);

x represents values of size (one independent variable); and

y represents the values of the socioeconomic status variable (the second independent variable).

Holding y constant, and differentiating z with respect to x, the relevant partial derivative is given by the equation:

$$f_x(z) = a + cy$$

(equation 2)

This equation can be used to calculate the effect on the dependent variable (z, achievement) for differing values of the variable held constant during partial differentiation (y, socioeconomic status).

Standardizing the partial derivative renders it as an "effect size," which is more easily interpretable than the unstandardized form[25].

---

[25]The derivation of effect size is elementary algebra and it proceeds as follows:

* Recall that a + cy = change in dependent variable (achievement) per change in size and, thus,
* a + cy represents a rate, call it dz/dx (i.e., change in achievement per change in size);
* to standardize, assume that the ES = some k, such that
* $(dz/dx) = k (\sigma_z / \sigma_x)$
  [that is, some constant times the ratio of the standard deviation for achievement and the standard deviation for size will equal the rate of change in achievement per change in size, i.e., the partial derivative], and therefore:
* $(dz/dx)(\sigma_x/\sigma_z) = k$
* and since (a +cy) =dz/dx, the untransformed rate,
* $k = (a + cy)(\sigma_x/\sigma_z)$, as given in the following discussion.

The total, standardized effect of school size on achievement (in standard deviation units) is given by the following formula:

$$\text{effect size} = (a + cy)(\sigma_x/\sigma_z) \qquad \text{(equation 3)}$$

This is the final form of the regression equations, and it represents the change in achievement (in standard deviation units) expected with change in size (also in standard deviation units) among cases with a particular SES.

APPENDIX B

SPSS regression output for those Table 4 and Table 5 equations with significant interaction terms, recomputed with pupil-teacher ratio as control

All tests, grade 4

Coefficients

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	61.670	2.563		24.060	.000
	PUPTCH94	-.290	.162	-.127	-1.786	.075
	centered sqrt_afl	-2.568	.243	-.549	-10.562	.000
	centerd ln_size	1.395	.628	.160	2.220	.027
	ctrses1*ctrsize	-.626	.168	-.196	-3.724	.000

a Dependent Variable: GR4MEAN

CTBS, grade 4

Coefficients

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	60.084	4.022		14.940	.000
	PUPTCH94	-2.376E-02	.268	-.011	-.089	.930
	centered sqrt_afl	-2.077	.384	-.425	-5.408	.000
	centerd ln_size	-1.873	1.128	-.195	-1.660	.100
	ctrses1*ctrsize	-1.588	.347	-.356	-4.573	.000

a Dependent Variable: GR4MEAN

ITBS, grade 4

Coefficients

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	61.529	3.631		16.944	.000
	PUPTCH94	-.303	.220	-.134	-1.374	.172
	centered sqrt_afl	-2.677	.324	-.665	-8.265	.000
	centerd ln_size	1.977	.768	.250	2.573	.011
	ctrses1*ctrsize	-1.102	.215	-.401	-5.130	.000

a Dependent Variable: GR4MEAN

Stanford, grade 4

Coefficients

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	58.079	5.953		9.756	.000
	PUPTCH94	-8.424E-02	.380	-.039	-.222	.825
	centered sqrt_afl	-2.383	.569	-.486	-4.186	.000
	centerd ln_size	-.739	1.792	-.092	-.412	.681
	ctrses1*ctrsize	.734	.395	.273	1.858	.068

a Dependent Variable: GR4MEAN

ITBS, grade 8

Coefficients

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	56.712	2.886		19.648	.000
	PUPTCH9-94	-7.965E-03	.207	-.005	-.039	.969
	centered sqrt_afl	-2.046	.472	-.546	-4.330	.000
	centerd ln_size	-.651	.669	-.127	-.973	.334
	ctrses1*ctrsize	-.838	.277	-.380	-3.025	.003

a Dependent Variable: GR8MEAN